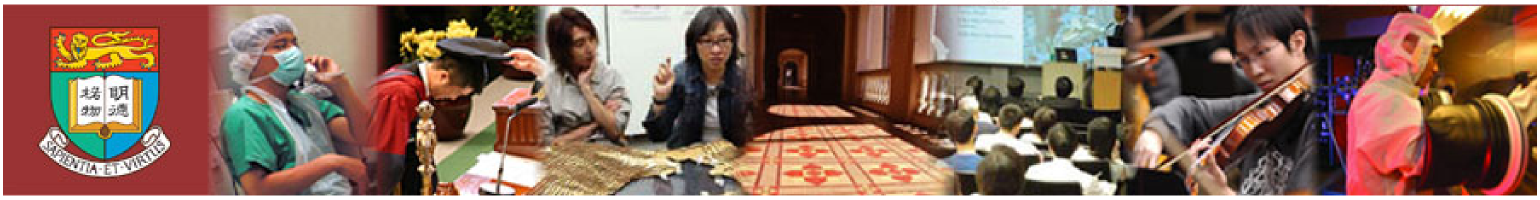


Title	Vocal tract dimensional characteristics of professional male and female singers with different types of singing voices
Author(s)	Tsoi, Kap-suen; 蔡甲算
Citation	
Issued Date	2009
URL	http://hdl.handle.net/10722/173686
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Vocal Tract Dimensional Characteristics of Professional Male and Female Singers with Different Types of Singing Voices

Tsoi Kap Suen

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2009

Vocal Tract Dimensional Characteristics of Professional Male and Female Singers with Different Types of Singing Voices

Tsoi Kap Suen

Abstract

This study investigated the relationship between the singers' singing voice classification and their vocal tract length, volume, and vowel formant frequencies. Five tenors, 5 baritones, 13 sopranos, and 5 mezzo-sopranos were included in this study. This study used the acoustic pharyngometer (ART) to measure the length and volume of the participants' vocal tract. The first three formant frequencies of vowels produced by the participants were analyzed with Praat. The results showed that higher singing voice types have higher formant frequencies. Baritones had greater vocal tract volume than tenors. These indicated that, besides vocal tract length, vocal tract volume may also affects the formant formation, and thus the singing voice classification.

Introduction

The vocal tract

The vocal tract is the cavity between the lips and the glottis. It consists of the oral cavity, the pharyngeal cavity, and the nasal cavity (Rod, 1995). The articulators, such as the lips, tongue, jaw, soft palate, and the larynx can alter the length and the volume of the vocal tract. For example, lowering the larynx while yawning results in the lengthening of the vocal tract. The vocal tract serves as the resonator in human voice production (Titze, 1994). The function of the vocal tract as the resonator of human voice production has been explained with the source-filter theory and the tube resonator model (Titze, 1994). When the air stream from the lungs passes through the vocal folds, the vocal folds vibrate. This vibration creates the sound source for voicing. This sound source is a complex wave that composes of the fundamental frequency (F0), which is determined by the frequency of the vibration of the vocal folds, and a large number of overtones, which are multiples of the fundamental frequency: the first overtone is one time higher than the fundamental frequency, and the second overtone is two times higher than the fundamental frequency (Titze, 1994). Depending of the length and the volume of the vocal tract, certain frequencies of the sound source are amplified. The amplified frequencies are referred as formant frequencies. By controlling the articulators, the shape of the vocal tract is altered, and different speech sounds can be produced. For example, Kent (1997) found that the positioning of the tongue mainly changes the first formant frequency (F1) and the second formant frequency (F2), while the third formant frequency (F3) and the fourth formant frequency (F4) were generally not affected. It was found that the higher the tongue position, the lower the F1; the more anterior the tongue position, the higher the F2. Without the amplification generated by the vocal tract, human voice resembles the sound created by a vibrating rubber band, and the pronunciations required for speech and singing cannot be achieved (Kent, 1997).

The singing voice classification in the singing community

Traditionally, vocal pedagogues divided singers into at least three main singing voice types in each gender: bass, baritone, and tenor for males, and alto, mezzo-soprano, and soprano for females (Titze, 1994). A certain type of singing voice possesses particular attributes, and can thus be distinguished from another type of singing voice. For vocal pedagogues, the voice range and the voice timbre are the two main factors in classifying a singer. The voice range refers to the range of the singing fundamental frequency that a singer can achieve. Titze (1994) provided the approximate singing fundamental frequencies of different singing voice types, which are summarized in table 1.

Table 1. Approximate ranges of singing fundamental frequency of different singing voice types (Titze,1994).

	Lower limit of fundamental frequency (Hz)	Upper limit of fundamental frequency (Hz)
Soprano	196.0	1174.7
Mezzo-soprano	164.8	880.0
Alto	146.8	587.3
Tenor	130.8	523.3
Baritone	98.0	392.0
Bass	82.4	329.6

This is obvious that the lower singing voice types, such as bass, produce lower fundamental frequency, while the higher singing voice types, such as soprano, produce higher fundamental frequency. However, it can be observed that the voice ranges of different types of singing voices overlap (Titze, 1994, see table 1). For example, the voice ranges of baritones and basses overlap with each other in more than 78% of their voice range (from 98.0 Hz to 329.0 Hz). However, this overlapping did not lead to the classification of those singers in the same type of singing voice.

Besides the voice range, vocal pedagogues also rely on the voice timbre during the singing voice classification. The generally accepted definition for voice timbre is as follow: when two tones are of the same pitch and loudness but are judged to be different, then the two tones are described to possess different timbres (ASNI, 1973). Even when two singers have the same voice range, they can still be classified in different singing voice type if their voices are of different timbre. This concept can be better understood if we consider singers as musical instruments. This is similar to the fact that a violin and a cello can play scales of the same frequency range, but one can easily identify the difference of the sound quality of the two instruments. However, this does not mean that a singer of a certain singing voice type sings with one particular timbre across the whole range of his/her voice. Erickson, Perry, & Handel (2001) suggested that singers of a certain singing voice type shared a similar pattern of “timbre transformation”. This “timbre transformation” is created by the altering of vocal registers in singing across a singer’s voice range. Sundberg (1987) defined a vocal register as a frequency range of phonation that consists of tones perceived to be produced in a similar way, and characterized with a similar voice timbre. Singers use a particular vocal register to produce a certain range of notes, or to achieve a particular voice timbre.

Researchers have been studying the physiology of the production of different vocal registers, and their corresponding frequency ranges (Hollien, 1974; Miller, 2000). Unfortunately, as Henrich (2006) pointed out, despite the effort of various researchers, the issue of vocal register was still controversial in the singing voice community, and its definition and labeling remained confusing. However, there are some generally agreed notions: 1) vocal registers are dissimilar in their voice timbre, 2) although vocal registers correspond to different frequency ranges, the frequency ranges of adjacent vocal registers overlap with various degree, 3) vocal registers can be roughly categorized as follow, with the order ascending in frequency range: the vocal fry register, or the pulse register (the lowest register); the modal register, or the chest register; the falsetto register, or the head register; and the whistle register, or the flute register (the highest register) (Bickel, 2008; Chapman, 2006; Nair, 1999; Sataloff, 2005).

With the above notions for the vocal register, the “timbre transformation” can be better understood. Professional singers are trained to use different vocal registers to accomplish the requirement of the “tessitura” of various pieces of song. Tessitura refers to the pitch range of a song or a singer (Bickel, 2008). In order to perform a song of a large tessitura, a singer may use different vocal registers, according to the pitches of the notes. For the singers in the same type of singing voice, their patterns of altering between different vocal registers across their voice range are very similar (Erickson, 2004). As the timbres of the vocal registers are dissimilar, the altering of vocal register will result in a “timbre transformation”. In the singing community, “timbre transformation” of singers is another factor that contributes to the singing voice classification (Bickel, 2008).

The singing voice classification in singing voice researches

Vocal pedagogues have been using the voice range and voice timbre to classify singers into different singing voice types. Whereas researchers have been interested in identifying the factors that create the differences between the singing voice types: what create the differences between the singing voice types and thus enables vocal pedagogues to differentiate them (Agren & Sundberg, 1976; Berndtsson & Sundberg, 1995, Cleveland, 1977; Dmitriev & Kiselev, 1979; Erickson, 2003, Erickson, 2004, Sundberg, 1973, Sundberg, 1979)? The study of Agren and Sundberg (1976) suggested the amplitude of the fundamental frequency as the main difference between the tenor voice and the alto voice. They compared the source spectra of five vowels (/e/, /i/, /a/, /o/, /u/) sung by two alto singers and two tenor singers. The source spectrum is a graphical illustration of the amplitude of voice across the frequency range (Nair, 1999). No consistent difference along the spectra was found between the two types of singing voice. However, it was found that the amplitude of the fundamental frequency in the alto voice was significantly higher than that of the tenor voice. This suggested the fundamental frequency as one of the contributing factors for the classification of different types of singing voice. However, as only four subjects were included in the study, and only the alto voice and tenor voice were compared, general conclusion could not be drawn.

The research of Cleveland (1977) came up with similar conclusion as Agren and Sundberg (1976). However, instead of the amplitude of the fundamental frequency, he suggested the pitch of the fundamental frequency as the main factor for the differentiation of various singing voices. In the study of Cleveland, eight professional singers were asked to sing five vowels (/e/, /i/, /a/, /o/, /u/) of four pitches (C3, F3, A3, E4). Vocal pedagogues were then asked to classify the vocalizations as voices produced by a tenor singer, a baritone singer, or a bass singer. The pitch, formant frequency, and source spectra were analyzed in order to find out the contributing factors in the classification of singing voice. The results showed that in the classification of singing voice, the pitch of the fundamental frequency served as the “primary cue”, whereas the formant frequencies were the “secondary cue”. It was found that the vocal pedagogues mainly depended on the pitch of the fundamental frequency when classifying the voice stimuli into a particular type of singing voice. However, when asked to classify voice stimuli of the same vowel and pitch sung by tenor singers, baritone singers, and bass singers, the vocal pedagogues had to rely on the “secondary cue”, the formant frequencies. It was because the fundamental frequency was very similar as the stimuli were of the same pitch. Cleveland found that the mean value of the first four formant frequencies in tenor singers was higher than that of baritone singers, while bass singers had the lowest mean value of the first four formant frequencies. As formant frequencies are determined by the vocal tract configurations, the findings of Cleveland suggested that the vocal tract characteristics might be a contributing factor of the differentiation of singing voices. However, this study also lacked statistical significance, as only eight participants were included. Therefore, the generalization of the findings was not solid.

Berndtsson & Sundberg (1995) seemed to agree with Cleveland (1977) on the importance of the vocal tract configuration for the singing voice classification. Their study suggested that the singer’s formant was relevant to the classification of singing voice. The singer’s formant, also referred as the “singing formant” and “ring”, was an important acoustic characteristic observed in Western opera singers (Vennard, 1967). Rather than the name “singer’s formant” may imply, the singer’s formant is actually not a formant. The singer’s formant is indeed the clustering of the F3, F4, and F5 at about 3000Hz, commonly observed in the spectrum of professional singers’ singing voice. Researchers have suggested different hypothesis of the physiology of the production of the singer’s formant, but no consensus has been reached (Detweiler & Detweilier, 1995; Sundberg, 1974). However, the common perspective they shared was that the production of the singer’s formant is related to the configuration of vocal tract. In the study of Berndtsson & Sundberg, they presented synthesized stimuli to six singing teachers and asked them to classify the voices as tenor, baritone, bass, or alto. The F3, F4, and F5 of the stimuli were manipulated to simulate the altering of the singer’s formant. The results showed that higher singer’s formant was associated with the classification of higher singing voice types. Considering their findings, and the general agreement that the singer’s formant is produced by the configuration of vocal tract, it becomes logical to suggest that the singing voice classification is also related to the vocal tract.

There seems to be generally accepted that the classification of singing voice is dependent on both the 1) fundamental frequency, which is mainly affected by the length of the vocal folds, and 2) the formant frequencies, which are suggested to be affected by the vocal tract. Furthermore, the relationship between the fundamental frequency and the vocal fold length is clear: the shorter the vocal folds, the higher the fundamental frequency, and vise versa. The higher singing voice types tend to possess higher fundamental frequencies. However, the relationship between the formant frequencies and the vocal tract, and their interaction with singing voice classification is still uncertain.

Dmitriev and Kiselev (1979) were the first to address the issue. They measured the vocal tract length of 20 Russian professional singers with lateral x-ray technology. In their research, the length of the vocal tract and the corresponding acoustic characteristics (the high and low singing formants) of 20 professional opera singers of different types of singing voice, namely bass, baritone, tenor, mezzo-soprano, soprano, and high soprano, were studied. It was found that each type of singing voice had a specific range of low singing formant frequency and high singing formant frequency (see table 2).

Table 2. Frequency of the low singing formant, frequency of the high singing formant, and the length of the vocal tract of different singing voice type (Dmitriev, & Kislev, 1979).

	Frequency of the low singing formant (Hz)	Frequency of the high singing formant (Hz)	Length of the vocal tract (cm)
High soprano	760-800	3100-3500	15.3-16.3
Soprano	700-760	2800-3100	16.8-18.5
Mezzo-soprano	540-600	2700-3000	16.7-18.3
Tenor	600-640	2700-2900	19.0-22.0
Baritone	540-600	2500-2700	21.5-24.0
Bass	450-540	2300-2500	23.3-25.0

Furthermore, the vocal tract length changed systematically according to the type of the singing voice: for female singers, sopranos had shorter vocal tract than mezzo-sopranos did; for male singers, tenors had the shortest vocal tract, followed by baritones, and basses (see table 2). It was suggested that the professional opera singers used a strictly fixed vocal tract configuration (a certain range of vocal tract length) for all vowels, and in the whole voice range of his/her type of singing voice. That is, for a professional singer of a certain type of singing voice, even the pitches and the articulations of his/her singing are changed significantly, his/her vocal tract length still remained almost constant. Therefore, the researchers came to the conclusion that the types of singing voice are correlated with the length of the vocal tract. However, as only the length of the vocal tract was measured, it was not know whether the volume of the vocal tract also interacted with the singing voice timbre, and thus the singing voice classification. Furthermore, there were disadvantages of using x-ray technology as a measure of human vocal tract. First, during the x-ray measure, both the participants and the examiners are exposed to radiation and they may suffer from the potential side effects. Secondly, the x-ray technology can only provide two-dimensional information of human vocal tract. Therefore, volumetric information cannot be obtained directly.

The ART Technology

In the current study, a noninvasive device for 3-dimential vocal tract parameters measure, the acoustic pharyngometer (ART, Eccovision Acoustic Pharyngometer; Sensormedics Corp., Yorba Linda, CA), was used to investigate the relationship between the dimensional characteristics (the length and volume) of vocal tract and singing voice classification. The ART makes use of the acoustic reflection (AR) technology, which was originally developed for a diagnostic system of the upper respiratory airway diseases (Xue & Hao, 2006). This technology has been widely used to measure the upper airway clinically and physiologically: nasal airway measure (Corey, Gungor, Nelson, Liu, & Fredberg, 1998), pharyngeal area measure (Huang, Sheng, Takahashi, Fukunaga, Toga, Takahashi, & Ohya, 1998), endotracheal tube positioning (Eckmann, Glassenberg, & Gavriety, 1996), and pharyngeal anatomy of patients with sleep apnea (Bradley, Brown, Grossman, Zamel, Martinex, & Phillipson, 1986). ART has also been used to measure the senile changes in the vocal tract (Xue & Hao, 2003; Xue, Jiang, Lin, Glassenberg, & Mueller, 1999). The ART utilizes acoustic reflection signals to generate graphical representation of the measurement of the cross-sectional area from the glottis to the lips. The probe tube (wave tube) generates an audible sound signal. This signal is transmitted into the cavity being examined through a coupler (mouthpiece), which is connected to the mouth of the participant. The acoustic pulse will be reflected partially when it encounters variations in the area of the cavity. The changes in the cross-sectional area of the cavity can be calculated by comparing the amplitude and temporal changes of the reflected acoustic pulse and the incident pulse. Therefore, the information of the length and the volume of the vocal tract can both be directly obtained, which makes the ART a more convenient means for vocal tract measure than x-ray technology. The process of ART measure is rapid and noninvasive. Furthermore, the accuracy of AR technology in delineating the changes of length and volume of the human vocal tract has been well validated with CT scan (D’Urzo, Rubinstein, Lawson, Vassal, Rebuck, & Slutsky, 1988; Min & Jiang, 1995) and MRI (Dang, Honda, & Suzuki, 1994; Corey, Gungor, Nelson, Fredberg, & Lai, 1997).

The aim of this study is to further investigate the relationship between the singing voice classification and vocal tract length, vocal tract volume, and formant frequencies. It is hypothesized that both the length and the volume of the vocal tract change systematically with different types of singing voice. That is, the length of the vocal tract decreases in the following order: baritone, tenor in male singers; mezzo-soprano, soprano in female singers. Whereas the volume of the vocal tract decreases in the following order: baritone, tenor in male singers; mezzo-soprano, soprano in female singers. The formant frequencies will decrease in the following order: tenor, baritone in male singers; soprano, mezzo-soprano in female singers. It is expected that the higher singing voice types possess shorter and smaller vocal tracts, and voices of higher formant frequencies.

Method

Participants

This study involved five tenors, five baritones, thirteen sopranos, and five mezzo-sopranos. All the singers were recruited form The Hong Kong Academy for Performing Arts (HKAPA). The type of singing voice and the professional quality of the participants were confirmed by the chief instructor of vocal studies of the HKAPA, who has more than thirty years of experience in opera performance. The descriptive statistics of the age, height, and weight of the participants are summarized in table 3.

Table 3. Descriptive statistics of the participants' age, height, and weight.

	Soprano (n=13)	Mezzo-soprano (n=5)	Tenor (n=5)	Baritone (n=5)
Age-mean (yr)	23.77	25.00	25.00	22.40
Age-SD (yr)	2.87	4.06	7.11	2.70
Age-range (yr)	19-28	20-29	17-36	20-27
Height-mean (m)	1.62	1.68	1.75	1.76
Height-SD (m)	0.06	0.05	0.07	0.06
Height-range (m)	1.50-1.68	1.62-1.73	1.69-1.87	1.68-1.84
Weight-mean (kg)	52.31	57.80	72.20	71.40
Weight-SD (kg)	6.37	2.95	14.32	9.63
Weight-range (kg)	42-63	53-60	59-95	60-86

All participants reported to have no history of craniofacial abnormalities, and no upper airway diseases at the time of testing. All the participants passed a pure-tone audiometry test bilaterally at 25dB level at the following frequencies: 250Hz, 500Hz, 1000Hz, 2000Hz, and 4000Hz.

Procedure

The length and the volume of the vocal tract of the participants were measured with the ART. Each participant was seated upright in a chair and breathed out air from his/her mouth to the wave tube through the mouthpiece. A corresponding area-distance curve that showed the cross-sectional area of the vocal tract as a function of the distance from the lips to the glottis was plotted. A total of three measurements were taken. Then the participants breathed out air from his/her nose, with the mouthpiece and the wave tube connected to his/her mouth. A total of four area-distance curves that showed the dimensional parameters of the participant's vocal tract were then obtained.

The participants' speech sounds and singing voices were recorded with the Praat software (Paul Boersma and David Weenink). During the recording, the participants seated upright with a 15 cm distance from a high-quality recorder, iSight (Apple Inc.), which was connected to an iBook G4 computer (Apple Inc.). The participants produced the vowels /a/, /i/, /u/, /e/, and /o/, and sang the song "happy birthday". The participants were asked to sing with a well-supported voice with modal register, without any adjustment (e.g. vibrato, portamento, staccato, etc.), and changes in vocal register. All recordings were made with a sampling rate of 48k Hz.

Analysis

One of the three area-distance curves was chosen for analysis with the consideration of the following two criteria: 1) the oral pharyngeal juncture (OPJ) of the mouth-breathing curve that matched the OPJ of the nose-breathing curve best, and 2) the curve that was most stable: with least fluctuations in magnitude caused by the changes in airflow. The selected area-distance curve was then divided into the following two regions according to the standard criteria outlined by the manufacturer: 1) the oral region: from the incisors to the anterior edge of the OPJ, and 2) the pharyngeal region: from the soft palate to the opening of the glottis. After the above procedures, the following six measurements will be obtained for each participant: oral volume, oral length, pharyngeal volume, pharyngeal length, vocal tract volume, and vocal tract length.

The participant's speech vowel formant frequencies and singing vowel formant frequencies were analyzed with the Praat software. Among the five spoken vowel sounds, the vowel /a/ sound was chosen for analysis. Whereas three sung vowel sounds were extracted for analysis: the vowel /i/ in the second syllable of the word “happy”, the vowel /ae/ in the first syllable of the word “happy”, and the vowel /u/ in the word “you”. The first three formant frequencies of the selected vowel sounds were obtained with the auto-tracking function of the Praat software.

Reliability of the measurements

As a measure for the interexaminer reliability, six participants were randomly selected for reanalysis in ART by the second investigator. The mean value of the six vocal tract parameters (oral volume, oral length, pharyngeal volume, pharyngeal length, vocal tract volume, and vocal tract length) obtained by the first investigator and the second investigator, their absolute difference, and the Spearman correlation coefficients are showed in table 4.

Table 4. Statistic data for interexaminer reliability of the vocal tract parameter measures with ART

Vocal tract parameters	Measure by first examiner	Measure by second examiner	Absolute difference	Spearman rho
OL (cm)	9.20	9.23	0.03	0.87
PL (cm)	7.70	7.38	0.32	0.88
VL (cm)	16.90	16.62	0.28	0.99
OV (ml)	47.63	49.92	2.29	0.92
PV (ml)	19.80	17.91	1.90	0.94
VV (ml)	67.43	67.82	0.39	0.93

Notes:

All the Spearman correlation coefficients have $p < .01$.

OL: oral length, PL: pharynx length, VL: vocal tract length,

OV: oral volume, PV: pharynx volume, VV: vocal tract volume

The correlations ($p < .01$) were significant for the measures of all the parameters, which indicated adequate reliability of the measurement.

Results

The aim of this study is to investigate the differences of the vocal tract characteristics between various types of singing voices. As there are fundamental physiological differences in the vocal tract characteristics between the two genders, the analysis will be complicated if the two genders are considered together. Therefore, in the analysis, data collected from the two genders will be compared separately in order to indicate the relationship between the vocal tract characteristics and various singing voice types.

Vocal tract length and volume

The means and the standard deviations of the six measured vocal tract parameters (oral volume, oral length, pharyngeal volume, pharyngeal length, vocal tract volume, and vocal tract length) are summarized in table 5.

Table 5. Means and standard deviations of six measured vocal tract parameters

Vocal tract parameters	Soprano		Mezzo-soprano		Tenor		Baritone	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
OL (cm)	8.80	0.54	8.90	0.46	8.96	0.51	9.50	0.41
PL (cm)	7.59	0.76	8.11	0.41	9.70	0.63	9.12	0.13
VL (cm)	16.39	0.84	17.01	0.43	18.66	0.23	18.62	0.38
OV (ml)	48.87	11.22	42.87	5.68	51.31	15.64	62.98	8.09
PV (ml)	18.50	4.44	21.10	2.09	29.28	5.34	28.92	11.21
VV (ml)	67.37	11.96	63.97	7.06	80.59	20.52	91.90	16.24

Notes:

OL: oral length, PL: pharynx length, VL: vocal tract length,

OV: oral volume, PV: pharynx volume, VV: vocal tract volume

For male singers, the vocal tract length and vocal tract volume of tenors and baritones were compared with the Mann-Whitney U test. No significant difference was found in the vocal tract length ($U = 11.00$, $z = -.314$, $p = .7533$; see figure 1), whereas significant difference was found in the vocal tract volume ($U = 2.00$, $z = -2.193$, $p = .0283$; see figure 2).

Figure 1. Means of vocal tract length of the four singing voice types

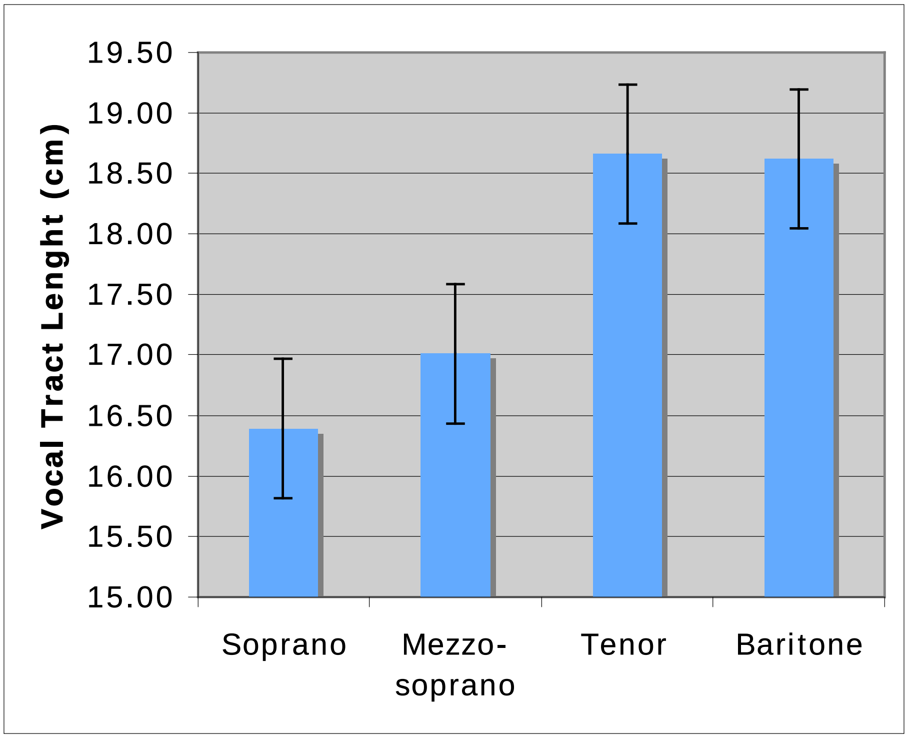
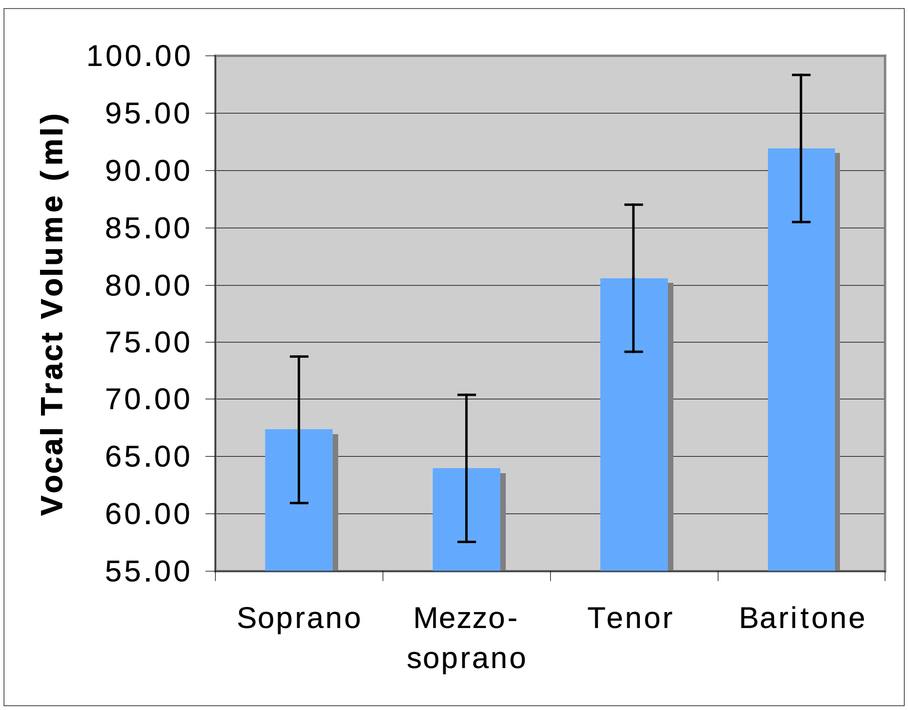


Figure 2. Means of vocal tract volume of the four singing voice types



Baritones were found to have significantly larger vocal tract volume than tenors. For female singers, the vocal tract length and vocal tract volume of sopranos and mezzo-sopranos were also compared with the Mann-Whitney U test. No significant differences were found in the vocal tract length ($U = 15.00$, $z = -1.729$, $p = .0837$, see figure 1), and the vocal tract volume ($U = 29.00$, $z = -.345$, $p = .7301$, see figure 2).

First three formant frequencies

The means and the standard deviations of the first three formant frequencies of the four selected vowels (/a/, /ae/, /i/, and /u/) are summarized in table 6.

Table 6. Means and standard deviations of the first formant frequencies of the four selected vowels

Formant frequencies	Soprano		Mezzo-soprano		Tenor		Baritone	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
F1 of /a/ (Hz)	779.12	170.89	678.45	52.72	667.09	82.77	593.42	86.24
F2 of /a/ (Hz)	1465.95	201.88	1162.90	41.29	1547.98	538.39	1166.28	138.67
F3 of /a/ (Hz)	2965.05	300.21	2883.76	118.23	2791.68	132.25	2662.49	170.98
F1 of /ae/ (Hz)	875.31	225.04	734.84	67.79	723.39	49.20	572.61	85.89
F2 of /ae/ (Hz)	1822.40	251.74	1616.43	296.91	1630.43	157.47	1464.96	56.38
F3 of /ae/ (Hz)	2788.58	321.11	2525.32	250.50	2651.13	105.52	2552.34	109.57
F1 of /i/ (Hz)	477.29	58.03	437.72	109.69	524.02	177.58	385.62	67.61
F2 of /i/ (Hz)	2167.74	173.67	1916.50	224.06	1943.15	227.01	1812.21	107.29
F3 of /i/ (Hz)	2816.63	241.88	2643.60	211.75	2573.22	113.69	2574.03	130.47
F1 of /u/ (Hz)	462.30	76.15	441.44	42.08	449.90	180.64	414.41	111.46
F2 of /u/ (Hz)	1749.56	249.95	1939.24	158.65	1588.26	725.49	1547.21	359.50
F3 of /u/ (Hz)	2786.65	210.54	2545.89	64.77	2654.54	143.47	2595.41	280.76

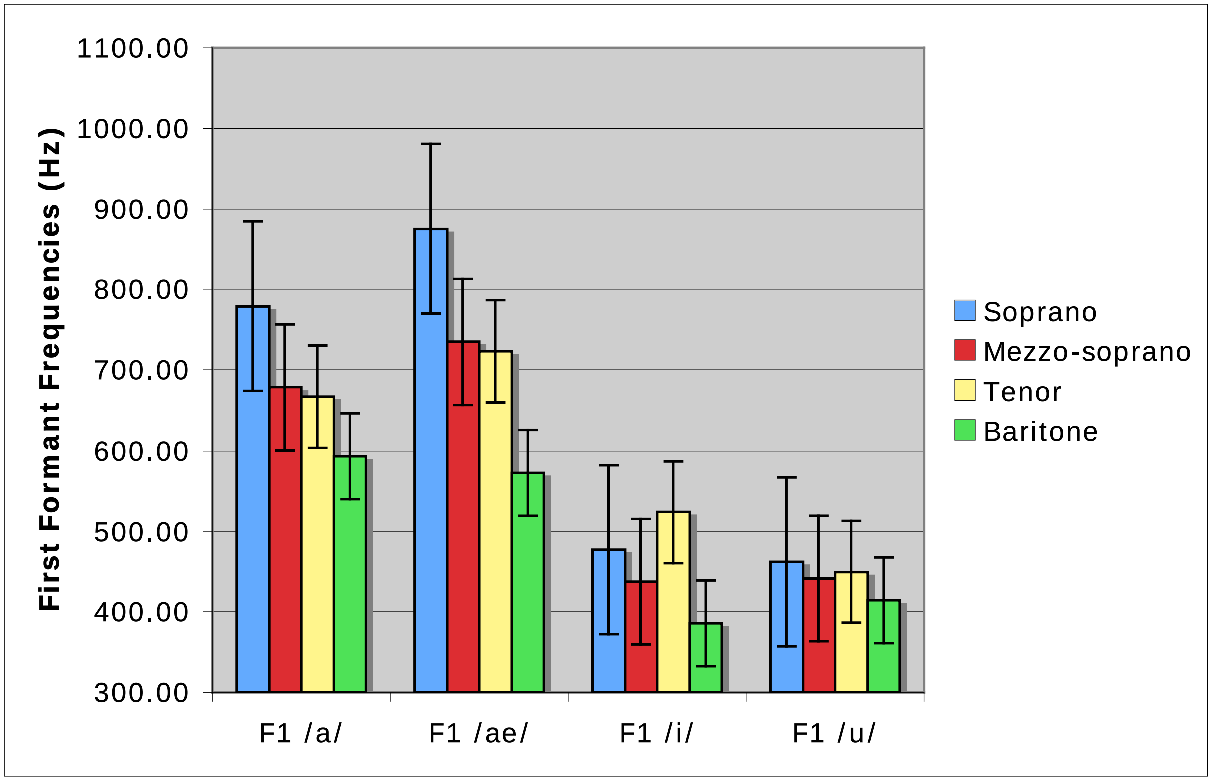
Notes:

F1: first formant frequency, F2: second formant frequency,

F3: third formant frequency

For male singers, the F1, F2, and F3 of the four selected vowel sounds produced by the tenors and baritones were compared with the Mann-Whitney U test. For F1, significant difference was found in the /ae/ sound (U = 2.00, z = 2.193, p = .0283), whereas no significant differences were found in the /a/ sound (U = 6.00, z = -1.358, p = .1745), the /i/ sound (U = 9.00, z = -.731, p = .4647), and the /u/ sound (U = 12.00, z = -.104, p = .9168). The F1 of the /ae/ sound produced by tenors was significantly higher than those of baritones (see figure 3).

Figure 3. Means of F1 of vowels /a/, /ae/, /i/, and /u/ of the four singing voice types



For F2, no significant differences were found in the /a/ sound (U = 5.00, z = -1.567, p = .1172), the /ae/ sound (U = 4.00, z = -1.776, p= 0.0758), the /i/ sound (U = 7.00, z = -1.149, p = .2506), and the /u/ sound (U = 12.00, z = -.104, p = .9168). For F3, no significant differences were found in the /a/ sound (U = 6.00, z = -1.358, p = .1745), the /ae/ sound (U = 5.00, z = -1.657, p = .1172), the /i/ sound (U = 12.00, z = -.104, p = .9168), and the /u/ sound (U = 10.00, z = -.522, p = .6015). The mean values of the first three formant frequencies of the four selected vowels sounds produced by the tenors and baritones (see table 7) were also compared with the Mann-Whitney U test.

Table 7. Means and standard deviations of averaged first three formant frequencies of the four selected vowels

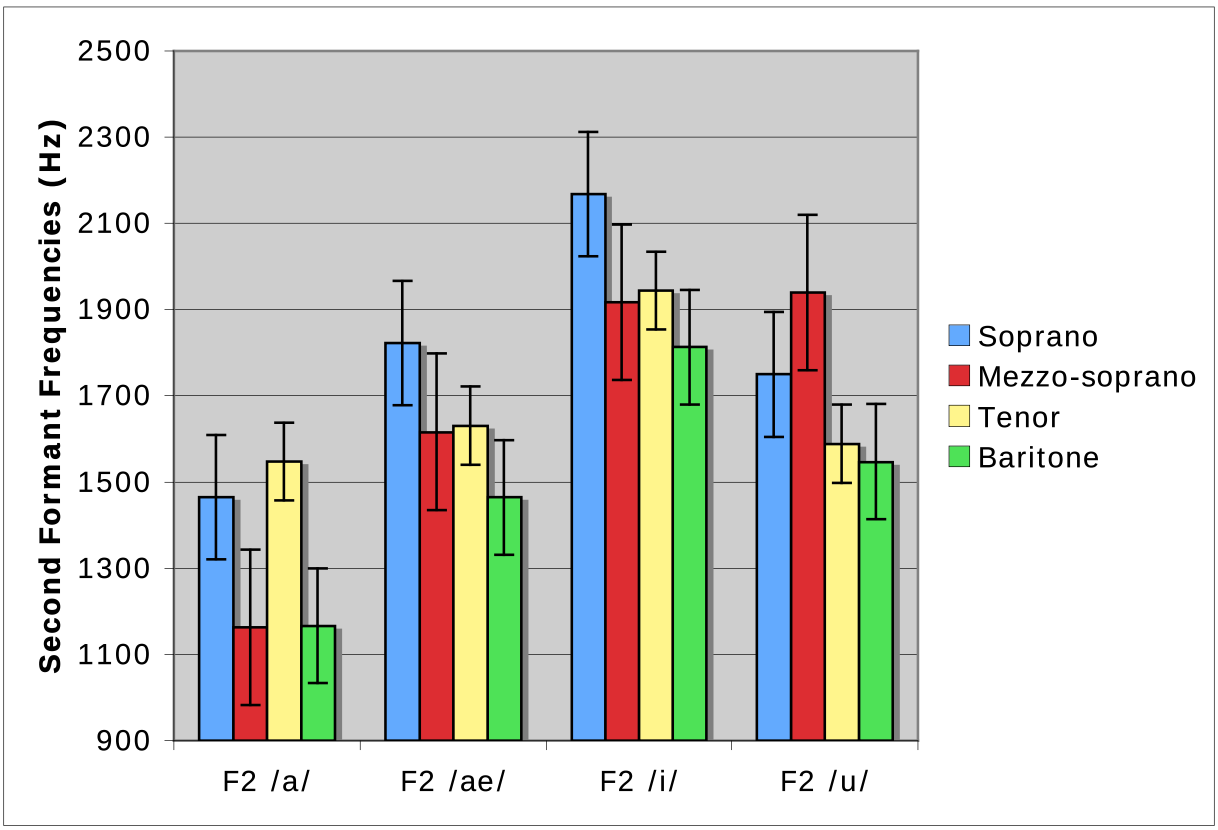
Averaged formant frequencies	Soprano		Mezzo-soprano		Tenor		Baritone	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/a/ (Hz)	1736.71	125.44	1575.04	21.09	1668.92	188.59	1474.07	76.26
/ae/ (Hz)	1828.76	201.43	1625.53	118.06	1668.32	81.28	1529.97	48.64
/i/ (Hz)	1820.55	135.63	1665.94	121.41	1680.13	135.72	1590.62	83.54
/u/ (Hz)	1666.17	129.48	1642.19	43.54	1564.24	343.06	1519.01	187.27

Significant difference was found in the /ae/ sound ($U = 2.00$, $z = -2.193$, $p = .0283$), whereas no significant differences were found in the /a/ sound ($U = 4.00$, $z = -1.776$, $p = .0758$), the /i/ sound ($U = 7.00$, $z = -1.149$, $p = .2506$), and the /u/ sound ($U = 12.00$, $z = -.104$, $p = .9168$). The mean value of the first three formant frequencies of the /ae/ sound was significantly higher than those of baritones (see figure 6). Although significant difference was only found in the /ae/ sound, the mean values of the first three formant frequencies of the four selected vowels produced by the tenors were generally higher than those of baritones (see figure 6).

For female singers, the F1, F2, and F3 of the four selected vowel sounds produced by the sopranos and mezzo-sopranos were compared with the Mann-Whitney U test. For F1, no significant differences were found in the /a/ sound ($U = 21.00$, $z = -1.134$, $p = .2570$), /ae/ sound ($U = 16.00$, $z = -1.626$, $p = .1039$), and /i/ sound ($U = 26.00$, $z = -.641$, $p = .5217$), and the /u/

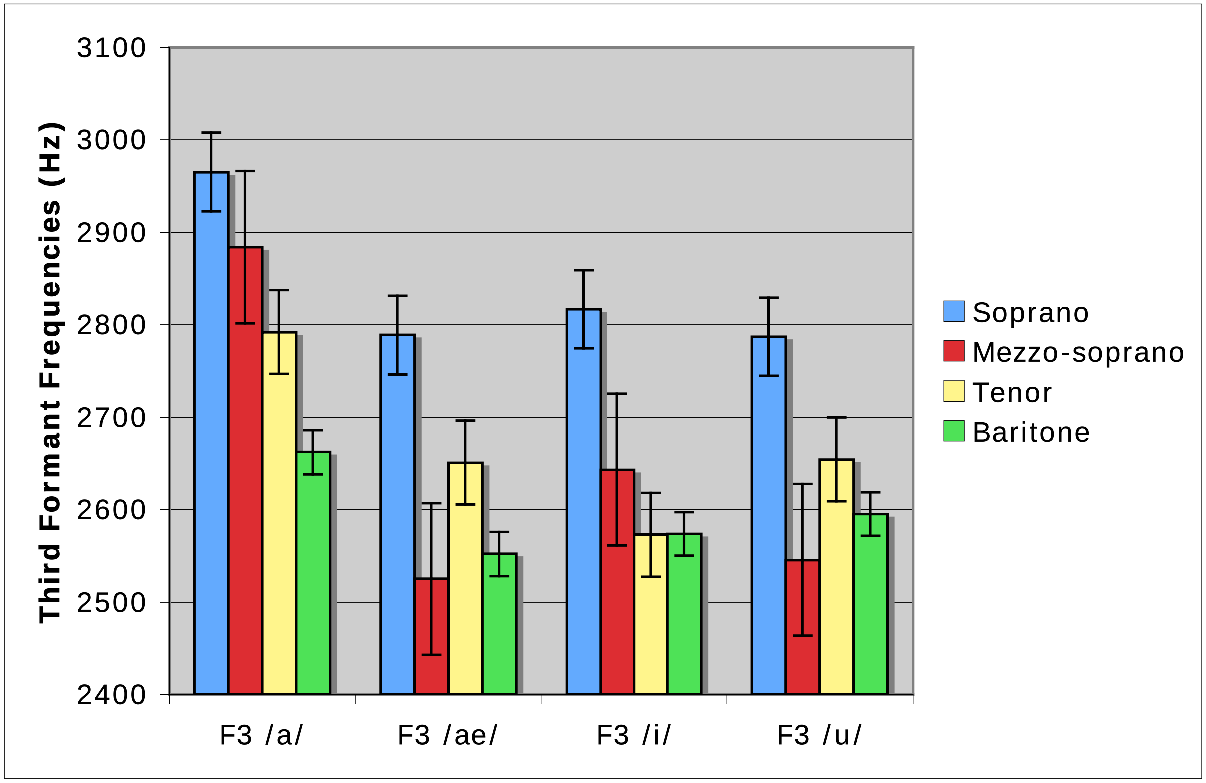
sound (U = 29.00, z = -.345, p = .7301). For F2, significant differences were found in the /a/ sound (U = 3.00, z = -2.908, p = .0036), and the /i/ sound (U = 12.00, z = -2.021, p = .0443), whereas no significant differences were found in the /ae/ sound (U = 18.00, z = -1.429, p = .1524), and the /u/ sound (U = 16.00, z = -1.626, p = .1039). The F2 of the /a/ sound, and the /i/ sound produced by sopranos were significantly highly than those of mezzo-sopranos (see figure 4).

Figure 4. Means of F2 of vowels /a/, /ae/, /i/, and /u/ of four singing voice types



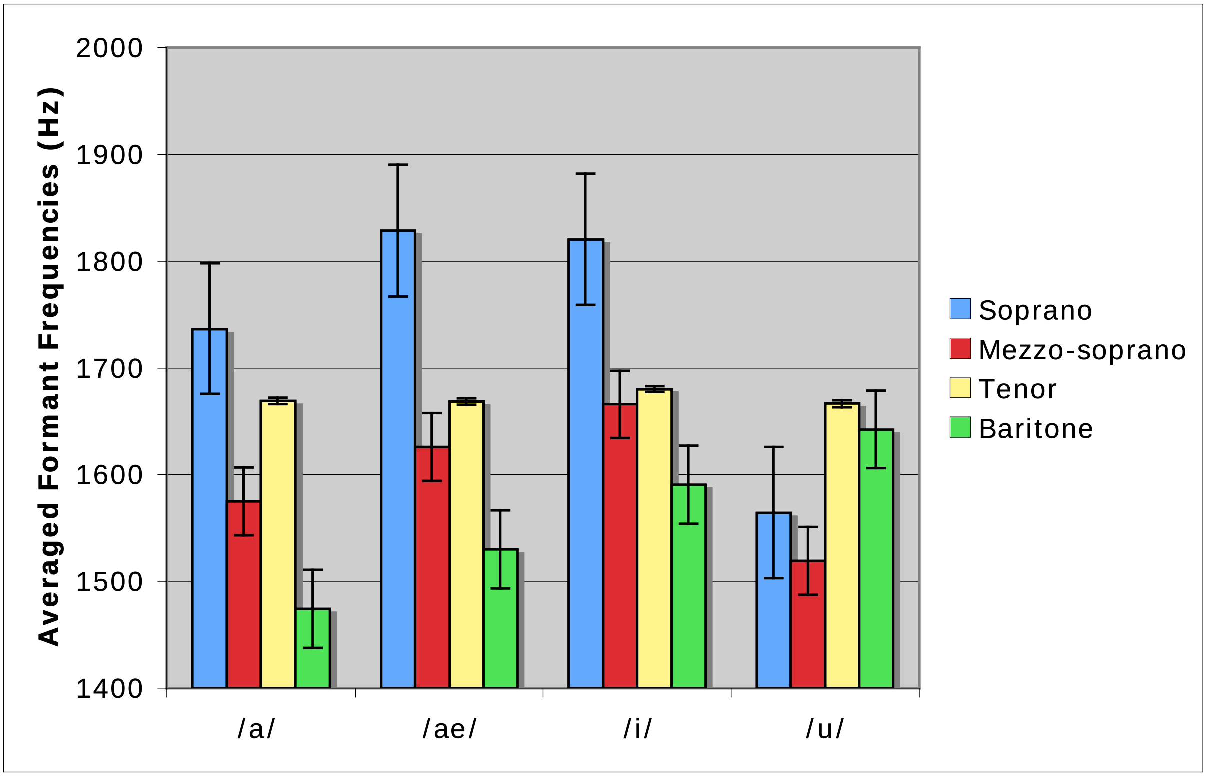
For F3, significant difference was found in the /u/ sound (U = 7.00, z = -2.514, p = .0120), whereas no significant difference was not found in the /a/ sound (U = 28.00, z = -.444, p = .6573), the /ae/ sound (U = 16.00, z = -1.626, p = .1039), and the /i/ sound (U = 21.00, z = -1.134, p = .2570). The F3 of the /u/ sound produced by sopranos were significantly higher than those of mezzo-sopranos (see figure 5).

Figure 5. Means of F3 of vowels /a/, /ae/, /i/, and /u/ of four singing voice types



The mean values of the first three formant frequencies of the four selected vowels sounds produced by the sopranos and mezzo-sopranos (see table 7) were also compared with the Mann-Whitney U test. Significant differences were found in the /a/ sound (U = 4.00, z = 2.809, p = .0050), and the /i/ sound (U = 12.00, z = -2.021, p = .0433), whereas no significant differences were found in the /ae/ sound (U = 14.00, z = -1.824, p = .0682), and the /u/ sound (U = 22.00, z = -1.035, p = .3007). The mean value of the first three formant frequencies of the /a/ sound, and the /i/ sound produced by sopranos were significantly higher than those of mezzo-sopranos (see figure 6).

Figure 6. Means of averaged formant frequencies of vowel /a/, /ae/, /i/, and /u/ of four singing voice types



Although significant difference was only found in the /a/ sound, and the /i/ sound, the mean values of the first three formant frequencies of the four selected vowels produced by the sopranos were generally higher than those of mezzo-sopranos (see figure 6).

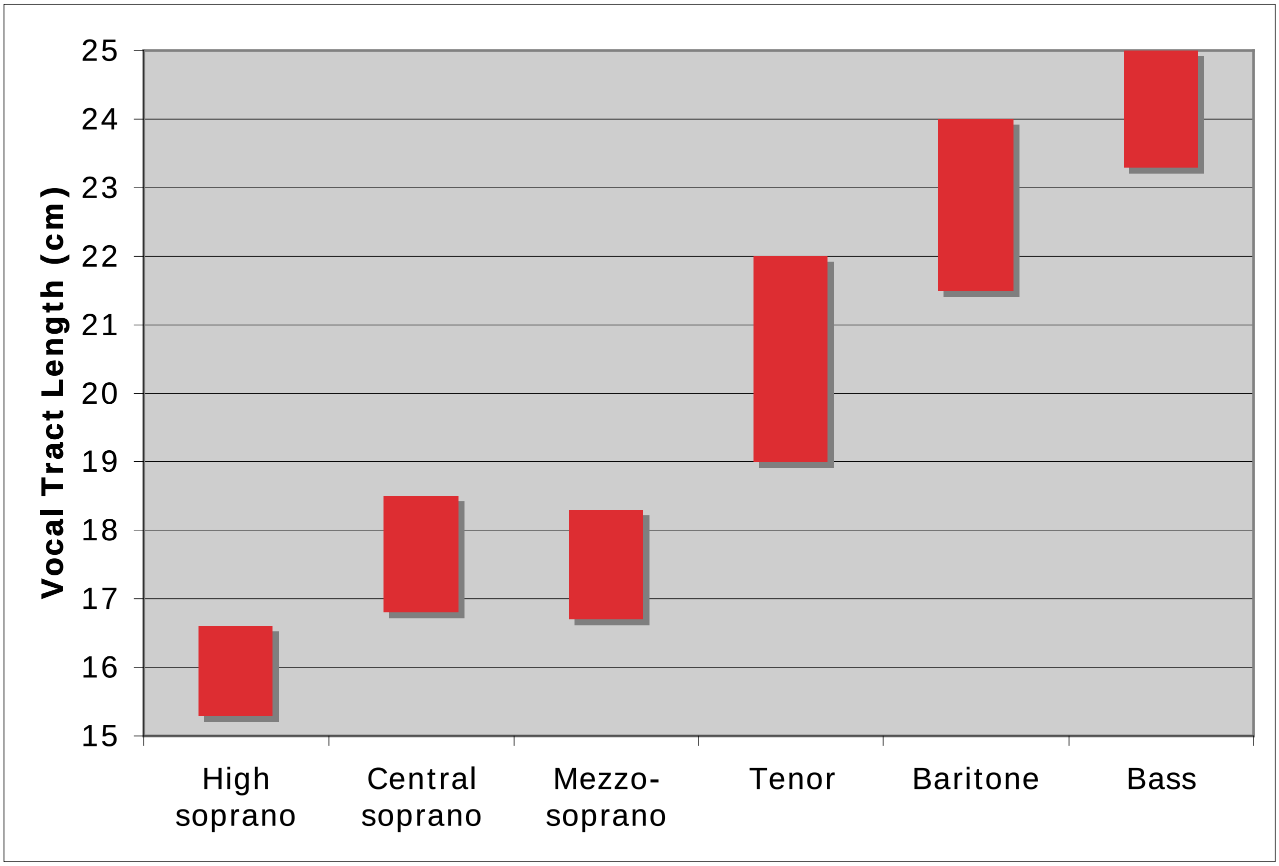
Discussion

There were three main findings in this study. First, in male singers, baritones possess a larger vocal tract volume than tenors do. Second, in male singers, the formant frequencies of tenors are generally higher than those of baritones. Third, in female singers, the formant frequencies of sopranos are generally higher than those of mezzo-sopranos.

In the study of Dmitriev & Kiselev (1979), they suggested that there was a correlation between the vocal tract length and the formant frequencies structure: the shorter the vocal tract, the higher the formant frequencies, which were correlated to higher singing voice types. Whereas the longer the vocal tract, the lower the formant frequencies structure, which were correlated to lower singing voice types. For male singers, Dmitriev & Kiselev found that baritone singers had longer vocal tract and lower formant frequencies than tenor singers. In the current study, the formant frequencies of baritone singers were also found to be lower than those of tenors. However, no significant difference was found between the vocal tract lengths of the two singing voice types. Nonetheless, the vocal tract volume of baritone singers was found to be significantly larger than that of tenors. For female singers, Dmitriev & Kiselev found that soprano singers had shorter vocal tract and higher formant frequencies than mezzo-sopranos. In this study, sopranos singers were also found to have higher formant frequencies. However, no significant difference was found between the vocal tract length of sopranos and mezzo-sopranos. Therefore, the findings of the current study have supported Dmitriev & Kiselev's work on their suggestion on the acoustical correlation between formant frequencies and singing voice types: higher formant frequencies are correlated with the higher singing voice types. However, the finding of the current study did not match with the finding of Dmitriev & Kiselev's on the relationship between the vocal tract length and the singing voice types. For male singers, significant difference was found in the vocal tract volume, but not in the vocal tract length of tenor singers and baritone singers. For female singers, no significant difference was found in both the vocal tract length and vocal tract volume.

Dmitriev & Kiselev (1979) suggested that tenor singers and baritone singers are different in their vocal tract length, which is directly related to the formation of the high and low formant frequencies. Tenors were shown to have vocal tract of shorter length, and singing formants of higher frequencies. However, their findings actually indicated some degree of overlapping in the range of vocal tract length of the two singing voice types (see figure 7).

Figure 7. Ranges of vocal tract length of six singing voice types (Dmitriev, & Kislev, 1979)



This implied that, even with very similar vocal tract lengths, two singers could still be classified in different singing voice types. Therefore, it seems logical to suggest that, besides the vocal tract length, some other factors may also affect the formation of formant frequencies and thus the characteristic voice timbre of a singer.

In this study, besides vocal tract length, the vocal tract volume of the singers was also investigated. For male singers, baritones were found to have significantly greater vocal tract volume, although the vocal tract lengths of the two singing voice types were found to be similar. Furthermore, baritones were also found to have significantly lower formant frequencies. By these observations, it is suggested that, besides vocal tract length, the volume of vocal tract may also affect the formation of formant frequencies, which in turn create the difference in timbre among the singing voice types. This argument becomes more feasible if the findings of Dmitriev & Kiselev (1979), and Erickson *et al.* (2001) are considered together. Dmitriev & Kiselev suggested that singers used a strictly fixed length of vocal tract across their whole voice ranges. Whereas Erickson *et al.* proposed a timbre transformation in singers as they proceeded from their low notes to high notes. If the vocal tract length is the only factor that creates the differences in timbre, the above two findings become contradictory. It is because, if the vocal tract length is the sole factor for the difference in timbre, a timbre transformation becomes impossible if singers use a fixed vocal tract length in their whole voice range. Therefore, other factors that interact with the voice timbre are bound to be present. The current study suggests vocal tract volume to be one of the possible factors that may affect the singing voice timbre. To further investigate the relationship between the vocal tract volume and the singing voice classification, study should be conducted with similar procedure, but with more singing voice types in each gender. For example, countertenors and basses can be included in male singers, whereas high-sopranos and altos can be involved in female singers. More representative findings may result when more singing voice types can be compared. The presence of contribution of vocal tract volume to singing voice classification may also be identified in female singers.

To further understand the relationship between the vocal tract characteristics and singing voice classification, studying the relationship between the timbre transformation and its corresponding changes in vocal tract may be useful. As Erickson *et al.* (2001) suggested, singers of a certain singing voice type share a similar pattern of timbre transformation. Therefore, it is expected that each singing voice type may demonstrate a characteristic timbre transformation, and thus a specific pattern of vocal tract configuration. To investigate this issue, ART may not be the best choice as phonation is not allowed during the measuring progress. Although professional singers are well trained in configuring their vocal tract for a specific voice timbre without phonation, the lack of simultaneous voice recording will negatively affect the validity of the study. The electromagnetic articulography (EMA) may be a better alternative. The EMA is a noninvasive device for the measurement of the movements of articulators in the vocal tract (Horn *et al.*, 1997; Kaburagi, Wakamiya, & Honda, 2005; Kaburagi & Honda, 2002). The dimensional characteristics of the vocal tract can be inferred with the sensors attached to the articulators. Furthermore, phonation will not interfere with the measure, so that simultaneous voice recording can be allowed.

References

Agren, K., & Sundberg, J. (1976). An acoustic comparison of alto and tenor voices. *Journal of Research in Singing, 1*, 26-32.

ASNI. (1973). *Psychoacoustical Terminology. S3.20.* New York: American National Standards Institute.

Berndtsson, G., & Sundberg, J. (1995). Perceptual significance of center frequency of the singer's formant. *Scaand J Logop Phoniatr, 20,* 35-41.

Bickel, J. E. (2008). *Vocal technique: a physiological approach for voice class and studio.* San Diego: Plural Publishing.

Bradley, T. D., Brown, I. G., Grossman, R. F., Zamel, N., Martinex, D., Phillipson, E. A., et al. (1986). Pharyngeal size in snorers, nonsnorers, and patients with obstructive sleep apnea. *New England Journal of Medicine*, *315*, 1327-1331.

Champman, J. L. (2006). *Singing and teaching singing: a holistic approach to classical voice.* San Diego: Plural Publishing.

Cleveland, T. F. (1977). Acoustic properties of voice timbre types and their influence on voice classification. *JASA*, *61,* 1622-1629.

Corey, J. P., Gungor, A., Nelson, R., Fredberg, J., & Lai, V. (1997). A comparison of the nasal cross-sectional areas and volumes obtained with acoustic rhinometry and magnetic resonance imaging. *Otolaryngology Head and Neck Surgery*. *117,* 349-354.

Corey, J. P., Gungor, A., Nelson, R., Liu, X., & Fredberg, J., (1998). Normative standards for nasal cross-sectional areas by race as measured by acoustic rhinometry. *Otolaryngology-Head and Neck Surgery*, *119,* 389-393.

Dang, J., Honda, K., & Suzuki, H. (1994). Morphological and acoustical analysis of the nasal and the paranasal cavities. *Journal of Acoustical Society of America. 96,* 2088-2100.

Detweiler, R., & Detweiler, G. (1995). *Investigation of the laryngeal system as the resonance source of the singer's formant: Data from a pedagogically heterogeneous population*, paper given at the 24th Annual Symposium Care of the Professional Voice, Philadelphia June 1995.

Dmitriev, L., & Kiselev, A. (1979). Relationship between the formant structure of different types of singing voices and the dimension of supraglottal cavities. *Folia Phoniat, 31,* 238-241.

D'Urzo, A. D., Rubinstein, I., Lawson, V. G., Vassal, K. P., Rebuck, A. S., et al. (1988). Comparison of glottic areas meauresd by scoustic reflection v.s. computerized tomography. *Journal of Applied Physiology. 127,* 221-226

Eckmann, D. M., Glassenberg, R., & Gavriety, N. (1996). Acoustic reflectometry and endotracheal intubation. *Anesthesia and Analgesia, 83,* 1084-1089.

Erickson, M. L., Perry, S., & Handel S. (2001). Discrimination functions: can they be used to classify singing voices? *Journal of Voice. 15(4),* 492-502.

Erickson, M. L. (2003). Dissimilarity and the classification of female singing voices, a preliminary study. *Journal of Voice. 17(2),* 195-206.

Erickson, M. L. (2004). The interaction of formant frequency and pitch in the perception of voice category and jaw opening in female singers. *Journal of Voice. 18(1),* 24-37.

Hirano, M., Vennard, W., & Ohala, J. (1970). Regulation of register, pitch, and intensity of voice. *Folia Phoniat. 22,* 1-20

Hollien, H. (1974). On vocal registers. *Journal of Phonetics, 2,* 125-143.

Horn, H., Goz, G., Bacher, M., Mullauer, M., Kretschmer, I., & Axmann-Krcmar, D. (1997). Reliability of electromagnetic articulography recording during speaking sequences. *The European Journal of Orthodontics. 19(6),* 647-655.

Huang, J., Shen, H., Takahashi, M., Fukunga, T., Toga, H., Takahashi, K., & Ohya, N. (1998). Pharyngeal cross-sectional area and pharyngeal compliance in normal male speakers and female speakers. *Respiration*, *65*, 458-468

Nair, G. (1999). *Voice tradition and technology: a state-of-the-art studio.* San Diego: Singular Publishing.

Kaburagi, T., & Honda, M. (2002). Electromagnetic articulograph based on a nonparametric representation of the magnetic field. *The Journal of the Acoustical Society of America. 111(3).* 1414-1421.

Kaburagi, T., Wakamiya, K., & Honda, M. (2005). Three-dimensional electromagnetic articulography: A measurement principle. *The Journal of the Acoustical Society of America. 118(1),* 428-443.

Kent, R. D. (1997). *The speech sciences.* San Diego: Singular Publishing Group.

Miller, D. G. (2000). *Registers in singing: empirical and systematic studies in the theory of the singing voice.* Groningen: University of Groningen.

Min, Y. G., & Jiang, Y. J. (1995). Measurements of cross-sectional area of the nasal cavity by acoustic rhinometry and CT scanning. *Laryngoscope*. *105*, 757-759

Rod, R. S. (1995). Anatomy of articulation and resonation. *Anatomy and physiology* (R.S. Rod, D. S. Trent, and T. Philip, Ed.). St. Louis: Mosby.

Sataloff, R. T. (2005). *Professional voice: the science and art of clinical care (3rd ed.).* San Diego: Plural Publishing.

Sundberg, J. (1973). The source spectrum of professional singing. *Folia Phoniatr, 25,* 71-90.

Sundberg, J. (1974). Articulatory interpretation of the 'singing formant'. *Journal of the Acoustical Society of America, 55,* 838-844.

Sundberg, J. (1977d). Studies on the soprano voice. *Journal of Research in Singing*, 1: 25-35.

Sundberg, J. (1979). Perception of singing voice. *STL-QPSR, 20,* 1-48.

Sundberg, J. (1987). *The science of the singing voice.* Dekalb: Northern Illinois University Press.

Titze, I. E. (1994). *Principles of voice production.* Englewood Cliffs, NJ: Prentice-Hall.

Vennard, W. (1967). *Singing - the mechanism and the technique*. New York: Carl Fisher, Inc.

Xue, S., & Hao, J. P. (2003). Gender differences in senile changes of human vocal tract and the acoustic correlates of speech production. *Journal of Speech, Language, and Hearing Research*, *46,* 689-701.

Xue, S., Jiang, J., Lin, E., Glassenberg, R., & Mueller, P. B. (1999). Age-related changes in human vocal tract configurations and the effects on the speakers' formant frequencies: a pilot study. *Logopedics Phonetics*, *13,* 245-252.

Xue, A., & Hao, J. G., (2006). Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry. *Journal of Voice*. *20(3)*, 391-400.

Appendix

Consent form for participants

Vocal tract dimensional characteristics of professional male and female singers with different types of singing voices

You are invited to participate in a research study conducted by Mr. Tsoi Kap Suen in the *Division of Speech and Hearing Sciences* at the University of Hong Kong.

PURPOSE OF THE STUDY
The aim of this study is to further investigate the relationship between different types of singing voice and their corresponding fundamental frequency, and vocal tract parameters, namely the length and volume of oral cavity, and pharyngeal cavity.

PROCEDURES
Your vocal tract dimensional characteristics will be measured with the acoustic pharyngometry (ART). You will also be asked to speak five vowel sounds, and sing the song “Happy Birthday”, which will be record with a computer. The procedure will last for about ten minutes.

POTENTIAL RISKS / DISCOMFORTS AND THEIR MINIMIZATION
The ART is a noninvasive device and will not generate any risks.

POTENTIAL BENEFITS
You may not have any benefit by participating in this study, but the data that you provide will contribute to the study of human vocal tract.

CONFIDENTIALITY
All the information obtained will be used for research only. Your name will be replaced by a code. Only the principal investigator, Mr. Tsoi Kap Suen, and his supervisor, Dr. Xue An, will have the access to the information.

PARTICIPATION AND WITHDRAWAL
Your participation is voluntary. This means that you can choose to stop at any time without negative consequences.

QUESTIONS AND CONCERNS
If you have any questions or concerns about the research, please feel free to contact Mr. Tsoi Kap Suen (6743-9102). If you have questions about your rights as a research participant, contact the Human Research Ethics Committee for Non-Clinical Faculties, HKU (2241-5267).

SIGNATURE

I ___________________________________ (Name of Participant)

understand the procedures described above and agree to participate in this study.

____
Signature of Participant

Date of Signature: ____________________

Date of Preparation: ____________________

HRECNCF Approval Expiration date: _____________________

Acknowledgement

This study is supported by Dr. Steve Xue, my supervisor, who provided me with excellent guidance; Mr. Brain Montgomery, the head of vocal studies of HKAPA, who helped in the arrangement on data collection; Ms. Regine Cheng and Ms. Ruby Ng, who assisted in the data collection; and the students of vocal studies of HKAPA, who spared their precious time during their preparation of annual performance for the data collection.